NASA/CR-1998- 207004

FINAL
IN-61-CR
15CIT
063071 January 30, 1998

Final Progress Report
NASA Grant NCC2-275

Research conducted with the support of NASA Grant NCC2-275 has been focused in the main on the development of fuzzy logic and soft computing methodologies and their applications to systems analysis and control, with emphasis on problem areas which are of relevance to NASA's missions.

One of the principal results of our research has been the development of a new methodology called Computing with Words (CW). Basically, in CW words drawn from a natural language are employed in place of numbers for computing and reasoning. There are two major imperatives for computing with words. First, computing with words is a necessity wen the available information is too imprecise to justify the use of numbers, and second, when there is a tolerance for imprecision which can be exploited to achieve tractability, robustness, low solution cost, and better rapport with reality. Exploitation of the tolerance for imprecision is an issue of central importance in CW.

In CW, the premises constitute the *initial data set* (IDS), which is assumed to consist of a collection of propositions expressed in a natural language. For purposes of computation, the propositions are expressed as canonical forms which serve to place in evidence the fuzzy constraints that are implicit in the premises. Then, the rules of inference in fuzzy logic are employed to propagate the constraints from premises to conclusions. As a final step, the induced constraints are re-translated into a natural language, resulting in what is referred to as the *terminal data set* (TDS).

A canonical form of a proposition p is expressed as

$$X \text{ isr } R$$

where X is the constrained variable, R is the constraining relation and isr is a variable copula in which the variable r defines the way in which R constrains X. In particular, a possibilistic constraint (r = blank) is expressed as

$$X \text{ is } R$$

and signifies that R is the possibility distribution of X. Similarly,

To: C.A.S.I.

$$X \text{ isp } P$$

in which P is a probability distribution, signifies that P is the probability distribution of X.

A canonical form of p serves to represent the meaning of p by placing in evidence the implicit constrained variable X and the implicit constraining relation R. The *depth* of p provides a measure of the difficulty of explicitating X and R.

One of the important objectives of our research was to develop a better understanding of the process of explicitation for fairly complex propositions exemplified by "It is unlikely that there will be a significant increase in the level of ozone in the stratosphere in the near future." Another important objective was to develop computationally efficient ways of applying the rules of inference in fuzzy logic to the propagation of constraints from premises to conclusions. In this context, our attention was focused on the so-called generalized extension principle which is expressed as

$$\frac{f(X) \text{ is R}}{g(X) \text{ is } g(f^{-1}(R))}$$

where X is a database variable, $f(X)$ is a given function of X which is constrained by R, and $g(X)$ represents the query. In general, employment of the generalized expansion principle reduced the computation of the terminal data set to the solution of a nonlinear program. In addition, we have explored the possibility of applying genetic computing to the solution of this problem.

Another important direction which was pursued relates to the development of fuzzy information granulation (fuzzy IG) as a basic methodology for the conception, design, construction and utilization of information/intelligent systems.

In essence, fuzzy information granulation may be viewed as a mode of generalization which may be applied to any concept, method or theory. Related to fuzzy IG are the following principal modes of generalization.

(a) fuzzification (f-generalization). In this mode of generalization, a crisp set is replaced by a fuzzy set.

(b) granulation (g-generalization). In this case, a set is partitioned into granules.

(c) randomization (r-generalization). In this case, a variable is replaced by a random variable.

(d) usualization (u-generalization). In this case, a proposition expressed as X is A is replaced with usually (X is A).

These and other modes of generalization may be employed in combination. A combination that is of particular importance is the conjunction of fuzzification and granulation. This combination plays a pivotal role in the theory of fuzzy information granulation and fuzzy logic, and is referred to as f.g-generalization (or f-granulation or fuzzy granulation).

As a mode of generalization, f.g-generalization may be applied to any concept, method or theory. In particular, in application to the basic concepts of variable, function and relation, f.g-generalization leads -- in fuzzy logic -- to the basic concepts of linguistic variable, fuzzy rule set and fuzzy graph. These concepts are unique to fuzzy logic and play a central role in its applications.

The distinctive concepts of f-generalization, g-generalization, r-generalization and f.g-generalization make a significant contribution to a better understanding of fuzzy logic and its relation to other methodologies for dealing with uncertainty and imprecision. In particular, crisp g-generalization of set theory and relational models of data lead to rough set theory. F-generalization of classical logic and set theory leads to multiple-valued logic, fuzzy logic in its narrow sense and parts of fuzzy set theory. But it is f.g-generalization that leads to fuzzy logic (FL) in its wide sense and underlies most of its applications. This is a key point that is frequently overlooked in discussions about fuzzy logic and its relation to other methodologies.

The point of departure in the theory of fuzzy information granulation is the concept of a generalized constraint. A generalized constraint on the values of a variable X is expressed as X isr R, where R is the constraining relation, isr is a variable copula and r is a discrete variable whose value defines the way in which R constrains X.

The principal types of constraints and the values of r which define them are the following:

1. *Possibilistic constraint*, $r = blank$. In this case, if R is a fuzzy set with membership function $\mu_R : U \to [0, 1]$, and X is a disjunctive (possibilistic) variable, that is, a variable which cannot be assigned two or more values in U simultaneously, then

$$X \; is \; R$$

means that R is the possibility distribution of X. More specifically,

$$X \; is \; R \to Poss\{X = u\} = \mu_R(u), u \in U.$$

2. *Veristic constraint*, $r = v$. In this case, if R is a fuzzy set with membership function μ_R and X is a conjunctive (veristic) variable, that is, a variable which can be assigned two or more values in U simultaneously, then

$$X \; isv \; R \to Ver\{X = u\} = \mu_R(u), u \in U$$

where $Ver\{X = u\}$ is the verity (truth value) of $X = u$.

It is important to observe that, in the case of a possibilistic constraint, the fuzzy set R plays the role of a possibility distribution, whereas in the case of a veristic constraint R plays the role of a verity distribution. What this implies is that, in general, any fuzzy -- and ipso facto any crisp -- set R admits of two different interpretations. More specifically, in the possibilistic interpretation the grades of membership are possibilities, while in the veristic interpretation the grades of membership are verities (truth values). Since in most cases constraints are possibilistic, the default assumption is that a fuzzy set plays the role of a possibility distribution.

3. *Probabilistic constraint*, $r = p$. In this case, $X\ isp\ R$ means that X is a random variable and R is the probability distribution (or density) of X. For example,

$$X\ isp\ N(m, \sigma^2)$$

means that X is a normally distributed random variable with mean m and variance σ^2. Similarly,

$$X\ isp\ (0.2\backslash a + 0.4\backslash b + 0.4\backslash c)$$

means that X takes the values a, b, c with respective probabilities 0.2, 0.4 and 0.4.

4. *Probability value constraint*, $r = \lambda$. In this case,

$$X\ is\lambda\ R$$

signifies that what is constrained is the probability of a specified event, $X\ is\ A$. More specifically,

$$X\ is\lambda\ R \rightarrow Prob\{X\ is\ A\}\ is\ R\ .$$

For example, if $A = small$ and $R = likely$, then

$$X\ is\lambda\ likely$$

means that

$$Prob\{X\ is\ small\}\ is\ likely\ .$$

5. *Random set constraint*, $r = rs$. In this case,

$$X\ isrs\ R$$

is a composite constraint which is a combination of probabilistic and possibilistic (or veristic) constraints. In a schematic form, a random set constraint may be represented as

$$Y\ isp\ P$$

$$\frac{(X,Y)\ is\ Q}{X\ isrs\ R}$$

or

$$Y\ isp\ P$$

$$\frac{(X,Y)\ isv\ Q}{X\ isrs\ R}$$

where Q is a joint possibilistic (or veristic) constraint on X and Y, and R is a random set, that is, a set-valued random variable. It is of interest to note that the Dempster-Shafer theory of evidence is in essence a theory of random set constraints.

A question that arises is: What purpose is served by having a large variety of constraints to choose from. A basic reason is that, in a general setting, information may be viewed as a constraint on a variable.

More generally, in the context of computing with words, a basic assumption is that a proposition, p, expressed in a natural language may be interpreted as a generalized constraint

$$p \rightarrow X\ isr\ R.$$

In this interpretation, $X\ isr\ R$ is the *canonical form* of p. The function of the canonical form is to place in evidence, i.e., explicitate, the implicit constraint which p represents.

In CW, the depth of explicitation of a proposition is a measure of the effort involved in explicitating p, that is, translating, p into its cannical form. In this sense, the proposition $X\ isr\ R$ is a surface constraint (depth = zero). A proposition such as "*X is small*" is shallow, whereas "it is not very likely that there will be a significant increase in the price of oil in the near future" is not.

The information conveyed by a proposition expressed in a natural language is, in general, too complex to admit of representation as a simple, crisp constraint. This is the main reason why in representing the meaning of a proposition expressed in a natural language we need a wide variety of constraints which are subsumed under the rubric of generalized constraints.

Beyond the development of the basic methodologies of computing with words and fuzzy information granulation, we have begun to explore their applications to knowledge representation, data mining, evolutionary computing and decision-making in an environment of imprecision, uncertainty and partial truth. In these applications, the available information is assumed to be granular in nature, reflecting the finite ability of the human mind and sensing devices to resolve detail. In particular, in the case of evolutionary computing and reinforcement learning, a basic assumption that is made is that evaluation functions are described as fuzzy graphs or, equivalently, by collections of fuzzy if-then rules.

A unifying idea which played a key role in our research is that in many real-world problems there is a built-in tolerance for imprecision. The effectiveness of the techniques of fuzzy logic derives from the fact that they exploit the tolerance for imprecision by simplifying the control strategies and enhancing their robustness. The importance of neural networks stems from their ability to learn from observations. The effectiveness of genetic algorithms results from their ability to optimize systems performance. In combination, these techniques form the core of soft computing.

Publications (partial list)

Hata, Y., Lee, M.A., and Yamato, K., "A Neuro-Fuzzy Computing Model of Human Pattern Generation," Proc. of the North American Fuzzy Information Processing Society (NAFIPS'96), Berkeley, CA, 1996.

Lee, M. A., and Esbensen, H., "The Design of Hybrid Fuzzy/Evolutionary Multiobjective Optimization Algorithms," 1995 IEEE/Nagoya University World Wiseperson Workshop Special Springer-Verlag Issue, 1996, Nagoya, Japan.

Lee, M. A., and Esbensen, H., "Multiobjective Optimization using Fuzzy/Evolutionary Algorithms," in Proc. of Int. Society for Computers and Their Applications (ISCA'96), San Francisco, CA, 1996, pp. 67-70.

Lee, M.A., and Esbensen, H., "Set Quality Measures for Characterizing Multiobjective Optimization Algorithm Behavior," Proc. of the North American Fuzzy Information Processing Society (NAFIPS'96), Berkeley, CA, 1996.

Lee, M.A., Takagi, H., "Hybrid Genetic-Fuzzy Systems for Intelligent Systems Design," in Genetic Algorithms and Soft Computing, Springer-Verlag, ed. F. Herrera and J.L. Verdegay, 1996.

Takagi, H., Lee, M.A., "Integration of Fuzzy Systems and Neural Networks and Fuzzy Systems and Genetic Algorithms," in Soft Computing Technologies, MIT Press, ed. L.C. Jan, 1996.

Zadeh, L.A., "Fuzzy Logic and the Calculi of Fuzzy Rules and Fuzzy Graphs: A Precis," IEEE Trans. on Fuzzy Systems, 1996.

Zadeh, L.A., "Fuzzy Logic = Computing with Words," IEEE Transactions on Fuzzy Systems, 2,103-111, 1996.

Zadeh, L.A., "Fuzzy Logic and the Calculi of Fuzzy Rules and Fuzzy Graphs," International Journal of Multiple Valued Logic, 1, 1-39, 1996.

Zadeh, L.A., "The Role of Fuzzy Logic and Soft Computing in the Conception, Design and Deployment of Intelligent Systems," BT Technical Journal, 14, No. 4, 32-36, 1996.

Zadeh, L.A., "Fuzzy Logic and Soft Computing," co-editor, World Scientific, 1996.

Zadeh, L.A., "Genetic Algorithms and Fuzzy Logic Systems," co-editor, World Scientific, 1997.

Zadeh, L.A., "Industrial Applications of Fuzzy Logic and Intelligent Systems," co-editor, IEEE Press, 1997.

Zadeh, L.A., "Applications of Fuzzy Logic," co-editor, Prentice-Hall, 1997.

Zadeh, L.A., "Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic," Fuzzy Sets and Systems, 90, 111-127, 1997.